

Insolubility of the Quantum Measurement Problem for Unsharp Observables

PAUL BUSCH

Department of Applied Mathematics, University of Hull

Hull, HU6 7RX, UK

E-mail: p.busch@maths.hull.ac.uk

ABNER SHIMONY

Departments of Physics and Philosophy, Boston University

Boston, MA 02215, USA

ABSTRACT

The quantum mechanical measurement problem is the difficulty of dealing with the indefiniteness of the pointer observable at the conclusion of a measurement process governed by unitary quantum dynamics. There has been hope to solve this problem by eliminating idealizations from the characterization of measurement. We state and prove two ‘insolubility theorems’ that disappoint this hope. In both the initial state of the apparatus is taken to be mixed rather than pure, and the correlation of the object observable and the pointer observable is allowed to be imperfect. In the *insolubility theorem for sharp observables*, which is only a modest extension of previous results, the object observable is taken to be an arbitrary projection valued measure. In the *insolubility theorem for unsharp observables*, which is essentially new, the object observable is taken to be a positive operator valued measure. Both theorems show that the measurement problem is not the consequence of neglecting the ever-present imperfections of actual measurements.

1. Introduction.

The quantum mechanical measurement problem consists of the difficulty of reconciling the occurrence of definite, objective ‘pointer’ readings with the unitarity and thus linearity of the time evolution of quantum states. The problem is posed simply if one considers the measurement of a discrete sharp observable $A = \sum_i a_i |\varphi_i\rangle\langle\varphi_i|$ and requires the following *calibration condition*: if the object system \mathcal{S} is in an eigenstate of A , say φ_k , then the state of the apparatus \mathcal{A} after the interaction with the object is an eigenstate of the pointer observable Z associated with a pointer reading z_k indicating that the value of A was a_k . Hence if the initial state of $\mathcal{S} + \mathcal{A}$ is $\varphi_k \otimes \phi$ and the measurement coupling is described by a unitary operator U , then the final state $U(\varphi_k \otimes \phi)$ must be an eigenstate of the observable $I \otimes Z$. But then the linearity of U entails that for any state φ of \mathcal{S} which is not an eigenstate of A , the resulting state of $\mathcal{S} + \mathcal{A}$ is a superposition of eigenstates of $I \otimes Z$. Certainly this does not correspond to a situation where the pointer observable is objectified.

It has sometimes been suggested that the quantum mechanical measurement problem is spurious, resulting from the excessively idealized characterization of the measurement process. One might conjecture that a more realistic characterization, taking into account the practical impossibility of preparing the macroscopic apparatus in a pure quantum state and acknowledging the possibility of imperfect correlation between the object observable and the pointer observable, would eliminate the measurement problem by assuring that the pointer observable is objectified. A series of papers, initiated by Wigner¹ and continued by various authors, including d’Espagnat², Fine³, and Shimony⁴, has thrown grave doubts on this conjecture in the important case when the object observable is a self-adjoint operator (equivalently, when it is represented by a spectral measure). In Sec. 2, the strongest result in this series, that of Ref. 4, is extended by taking the observable to be an arbitrary projection valued measure. We shall call the extended theorem the *insolubility theorem for sharp observables*.

The primary purpose of this paper is to consider the consequences of removing one more idealization in the usual treatments of measurement: viz., to replace the assumption that the object observables are sharp by the more realistic assumption that they are (or may be) unsharp. In Section 3 the measurement process of unsharp observables will be

characterized mathematically. We shall then demonstrate the main result of this paper: the *insolubility theorem for unsharp observables*, where an unsharp observable is represented as a positive operator valued measure.

2. Insolubility theorem for sharp observables.

In standard formulations of quantum mechanics, the *pure states* of a system are identified with rays (one-dimensional subspaces) of a Hilbert space \mathcal{H} , or equivalently, with projection operators onto rays. General *states* (including both *mixed* and *pure* states) are represented by density or state operators on \mathcal{H} . For any unit vector $\varphi \in \mathcal{H}$, the corresponding projection shall be denoted P_φ . *Observables* of the system are identified with self-adjoint operators on \mathcal{H} . (We use the terms *states* and *observables* both for the physical entities and the mathematical objects representing them.) By the classical spectral theorem of von Neumann, a self-adjoint operator A can be expressed in terms of a family of projection operators. In modern mathematical locution, with any self-adjoint operator A there is associated a unique projection valued measure, its spectral measure E^A . This assertion makes it possible to recover the expectation values $\langle A \rangle_\varphi$ over the probability measures $X \mapsto \langle \varphi | E^A(X) \varphi \rangle$ (where X runs through the (Borel) subsets of \mathbf{R}).

A first extension of the set of observables is obtained by admitting more general *projection valued (PV) measures*, defined with respect to a measurable space (Ω, Σ) , where Ω is a set and Σ is a σ -algebra of subsets of Ω : i.e., a PV measure is a map E from Σ into the lattice of projections such that $E(\emptyset) = O$, $E(\Omega) = I$ (O, I denoting the null and unit operators in \mathcal{H} , respectively), and $E(\cup_i X_i) = \sum_i E(X_i)$ for any countable collection of mutually disjoint sets $X_i \in \Sigma$. These conditions ensure that for any state operator T of \mathcal{S} , the map $X \mapsto \text{tr}[TE(X)] =: p_T^E(X)$ is a probability measure on (Ω, Σ) . Since Ω concerns the set of possible values of the physical quantity represented by E , (Ω, Σ) is called the *value space* of that observable. The case of a spectral measure is recovered by choosing the real line for Ω and the Borel sets for Σ . The introduction of more general value spaces (Ω, Σ) and of observables as PV measures on them proves convenient for a variety of purposes, such as, for example, the description of joint measurements of several commuting observables. Henceforth, we shall use the term *sharp observable* for an observable represented by

a general PV measure,⁵ in anticipation of a further generalization in Section 3 to unsharp observables.

In the remainder of this paper we shall take \mathcal{H} and \mathcal{H}_a to be the Hilbert spaces associated respectively with the object system \mathcal{S} and the apparatus \mathcal{A} . In the spirit of Ref. 5 we shall call the quadruple $\langle \mathcal{H}_a, Z, T_a, U \rangle$ a *measurement scheme* for \mathcal{S} if Z is the sharp pointer observable of \mathcal{A} , T_a an initial state of \mathcal{A} , and U a unitary operator on $\mathcal{H} \otimes \mathcal{H}_a$, the tensor product space associated with $\mathcal{S} + \mathcal{A}$. Note that the concept of a measurement scheme does not explicitly refer to an observable of the object \mathcal{S} , but the following useful concepts do so.

DEFINITION 1. Let E be a sharp observable of a system \mathcal{S} . Two state operators T and T' are E -distinguishable if and only if $p_T^E \neq p_{T'}^E$.

DEFINITION 2. A measurement scheme $\langle \mathcal{H}_a, Z, T_a, U \rangle$ for \mathcal{S} is a *discrimination* of the sharp observable E if and only if the E -distinguishability of any two states T, T' of \mathcal{S} implies the $I \otimes Z$ -distinguishability of $U(T \otimes T_a)U^{-1}$ and $U(T' \otimes T_a)U^{-1}$.

DEFINITION 3. Given a sharp observable E , a measurement scheme $\langle \mathcal{H}_a, Z, T_a, U \rangle$ is an E -measurement if and only if the PV measure Z is defined on the value space of E and the probability reproducibility condition is fulfilled: for all states T of \mathcal{S} and all $X \in \Sigma$,

$$p_T^E(X) = p_{U(T \otimes T_a)U^{-1}}^{I \otimes Z}(X) \quad (1)$$

Definitions 1 and 2 were essentially given by Fine³, but with E restricted to a spectral measure (equivalently, to a self-adjoint operator). We have departed from Fine's terminology, however, by using the term *E-discrimination* in Definition 2 when he uses the term *measurement*. E -discrimination as defined seems to us the weakest, hence the most general, condition connecting a measurement scheme with an observable, while intuitively measurement is a more stringent and special case of discrimination. Furthermore, Definition 3 singles out a unique observable as the one measured by a given measurement scheme, whereas Definition 2 does not do so.

In order to formulate the insolubility theorem for sharp observables in a way that applies to general PV measures, including continuous observables, we need the concept of

a *reading scale* for the pointer observable Z as a partition of its value space (Ω, Σ) , that is, a sequence of subsets $(X_i) \subseteq \Sigma$ such that $X_i \cap X_j = \emptyset$ if $i \neq j$ and $\cup X_i = \Omega$. Furthermore, it is necessary to exclude trivial PV measures from the discussion: a PV measure on (Ω, Σ) is *trivial* if there is a point $\omega \in \Omega$ such that $E(\{\omega\}) = I$. Equivalently, a PV measure is non-trivial if there are at least two E -distinguishable states.

INSOLUBILITY THEOREM FOR SHARP OBSERVABLES. *Let E be a non-trivial PV measure. There is no E -discrimination (and hence no E -measurement) $\langle \mathcal{H}_a, Z, T_a, U \rangle$ such that for all initial \mathcal{S} states T the $\mathcal{S} + \mathcal{A}$ states $U(T \otimes T_a)U^{-1}$ are mixtures of eigenstates of the projections $I \otimes Z(X_i)$ for a given reading scale (X_i) .*

This theorem was essentially formulated by Fine³ for the important case in which E is a discrete spectral measure. The proof which he gives, however, does not establish the theorem as he asserted it. As Brown⁶ points out, Fine's proof does establish a weaker theorem. This matter is discussed in Fine⁷ and Shimony⁸ (and see also the recent discussion of Stein⁹). Our proof of the insolubility theorem will follow (but with a generalization to an arbitrary PV measure) the procedure of Shimony⁴, who begins by proving an auxiliary theorem.

INHERITANCE OF SUPERPOSITIONS THEOREM. *Hypotheses:*

(i) φ_1, φ_2 are normalized orthogonal vectors of \mathcal{H} ; $\{F_m\}$ is a (finite or countable) family of mutually orthogonal projections of \mathcal{H}_a ; U is a unitary operator on $\mathcal{H} \otimes \mathcal{H}_a$, and T_a is a state operator on \mathcal{H}_a .

(ii) For some value of n

$$\text{tr}[I \otimes F_n U(P_{\varphi_1} \otimes T_a)U^{-1}] \neq \text{tr}[I \otimes F_n U(P_{\varphi_2} \otimes T_a)U^{-1}] \quad (2)$$

(iii) There exist orthonormal sets $\{\xi_{nr}^1\}$, $\{\xi_{nr}^2\}$ and sets of positive numbers $\{b_{nr}^1\}$, $\{b_{nr}^2\}$ such that $\xi_{nr}^j \in \mathcal{H} \otimes F_n(\mathcal{H}_a)$ for $j = 1, 2$ and

$$U(P_{\varphi_j} \otimes T_a)U^{-1} = \sum_{n,r} b_{nr}^j P_{\xi_{nr}^j} \quad (3)$$

Conclusion: If $\varphi \in \mathcal{H}$ is defined as $c_1\varphi_1 + c_2\varphi_2$, with both c_1 and c_2 nonzero, then there exists no orthonormal set $\{\psi_{nr}\}$ with $\psi_{nr} \in \mathcal{H} \otimes F_n(\mathcal{H}_a)$ and no positive coefficients $\{b_{nr}\}$

such that $\sum_{n,r} b_{nr} = 1$ and

$$U(P_\varphi \otimes T_a)U^{-1} = \sum_{n,r} b_{nr} P_{\psi_{nr}} \quad (4)$$

We do not reproduce the lengthy proof of this theorem here.⁴ Hypothesis (ii) is a minimal requirement of information transfer in the spirit of the concept of discrimination. Condition (iii) corresponds to the idea that a measurement coupling should lead to a final state which is a mixture of pointer eigenstates. Then the conclusion states that these two conditions cannot be reconciled with each other. It is not difficult to realize (4) if (ii) is given up. Let U be of the form $V \otimes V_a$, where V, V_a are unitary operators and V_a commutes with all F_n . Further let T_a be a mixture of F_n -eigenstates. Then $U(P_\varphi \otimes T_a)U^{-1} = VP_\varphi V^{-1} \otimes V_a T_a V_a^{-1}$ is a mixture of $I \otimes F_n$ -eigenstates, in fulfilment of (4). But the probabilities $\text{tr}[I \otimes F_n U(P_\varphi \otimes T_a)U^{-1}] = \text{tr}[F_n V_a T_a V_a^{-1}] = \text{tr}[F_n T_a]$ are independent of φ so that condition (ii) is violated.

We are now ready to prove the insolubility theorem for sharp observables. Suppose φ_1, φ_2 are two normalized eigenvectors of the projections $E(X), E(Y)$, respectively, with disjoint sets X, Y : i.e., $E(X)\varphi_1 = \varphi_1, E(Y)\varphi_2 = \varphi_2$. Then φ_1 and φ_2 are mutually orthogonal and indeed E -distinguishable. Suppose $U(P_{\varphi_1} \otimes T_a)U^{-1}$ and $U(P_{\varphi_2} \otimes T_a)U^{-1}$ are $I \otimes Z$ -distinguishable and are both expressible as mixtures of eigenstates of the projections $I \otimes Z(X_i)$. Then the hypotheses of the inheritance of superpositions theorem are satisfied. Therefore $U(P_\varphi \otimes T_a)U^{-1}$ is not a mixture of eigenstates of the $I \otimes Z(X_i)$, if φ is the superposition $c_1\varphi_1 + c_2\varphi_2$ with nonzero coefficients c_1, c_2 . On the other hand, if vectors φ_1, φ_2 with the assumed properties do not exist, then evidently the claim of the insolubility theorem would also hold.

As long as we restrict our attention to sharp observables, it is hard to envisage further strengthenings of this no-go verdict in the sense of extensions to more realistic measurement situations.

3. Insolubility theorem for unsharp observables.

So far the inevitable imperfections of actual measurements have been acknowledged by weakening the correlation between sharp object observables and pointer observables. Indeed, the concept of an E -discrimination in Definition 2 is an extreme expression of such

weakening. Another way of being realistic about measurements and acknowledging their imperfections is to take the object observables themselves to be unsharp.¹⁰ Therefore we consider now the case of a general observable, represented as a positive operator valued measure E on a measurable space (Ω, Σ) , its value space.

The map $E : X \mapsto E(X)$ is a *positive operator valued (POV) measure* if the following conditions are satisfied: for each $X \in \Sigma$, $E(X)$ is an operator on the underlying Hilbert space \mathcal{H} such that $O \leq E(X) \leq I$ [the ordering being in the sense of expectation values; i.e., $A \leq B$ if and only if $\langle \varphi | A \varphi \rangle \leq \langle \varphi | B \varphi \rangle$ for all $\varphi \in \mathcal{H}$]; moreover, $E(\emptyset) = O$ and $E(\Omega) = I$, and $E(\cup_i X_i) = \sum_i E(X_i)$ for any countable pairwise disjoint family $\{X_i\} \subseteq \Sigma$. These properties of E ensure that the map $p_T^E : X \mapsto p_T^E(X) := \text{tr}[TE(X)]$ is a probability measure for each state T . The special case of PV measures is recovered if the additional property of idempotency, $E(X)^2 = E(X)$, is stipulated.

Note that the idempotency condition can be written as $E(X)E(\Omega \setminus X) = O$, where $\Omega \setminus X$ is the complement of X so that $E(\Omega \setminus X) = I - E(X)$. It follows immediately that a POV measure is a sharp observable if and only if for any two disjoint sets X, Y the operators $E(X), E(Y)$ satisfy $E(X)E(Y) = O$. Such projections are orthogonal to each other in the sense that their ranges are mutually orthogonal subspaces. By contrast, for all other POV measures there will be sets X such that $E(X)$ and $E(\Omega \setminus X)$ are non-orthogonal. Such POV measures shall be called *unsharp observables*.¹¹

We now can generalize the concepts given in Definitions 1, 2 and 3. We use the term measurement scheme for $\langle \mathcal{H}_a, Z, T_a, U \rangle$ as given in Sec. 2. Note that the pointer Z is a PV measure even though we are generalizing the notion of an observable for the object system.

DEFINITION 1'. Let E be a POV measure associated with a system \mathcal{S} . Two state operators T and T' are E -distinguishable if and only if $p_T^E \neq p_{T'}^E$.

DEFINITION 2'. A measurement scheme $\langle \mathcal{H}_a, Z, T_a, U \rangle$ for \mathcal{S} is a *discrimination* of the POV measure E if and only if the E -distinguishability of any two states T, T' of \mathcal{S} implies the $I \otimes Z$ -distinguishability of $U(T \otimes T_a)U^{-1}$ and $U(T' \otimes T_a)U^{-1}$.

DEFINITION 3'. Given a POV measure E , a measurement scheme $\langle \mathcal{H}_a, Z, T_a, U \rangle$ is an E -measurement if and only if the PV measure Z is defined on the value space of E and

the probability reproducibility condition is fulfilled: for all states T of \mathcal{S} and all $X \in \Sigma$,

$$p_T^E(X) = p_{U(T \otimes T_a)U^{-1}}^{I \otimes Z}(X) \quad (5)$$

As mentioned in Section 2, one can read this definition in the opposite direction in the sense that any measurement scheme $\langle \mathcal{H}_a, Z, T_a, U \rangle$ induces, via Eq. (4), a unique observable E as the measured one. As an example, the probabilities $\text{tr}[I \otimes F_n U(P_\varphi \otimes T_a)U^{-1}]$ occurring in Eq. (2) can be expressed as $\langle \varphi | E_n \varphi \rangle$, defining thereby a set of positive operators E_n which constitute the measured observable given a pointer observable Z constituted by the projections F_n .

In the following only non-trivial observables E will be considered. E is *trivial* if and only if it is of the form $E : X \mapsto E(X) := \lambda(X)I$ for some probability measure λ . Equivalently, E is *non-trivial* if and only if there are at least two states which are E -distinguishable. Note that hypothesis (ii) of the inheritance of superpositions theorem ensures that the measured observable induced by the measurement scheme presented there is a non-trivial observable.

We now have all concepts in hand for stating the generalization of the insolubility theorem for unsharp observables.

INSOLUBILITY THEOREM FOR UNSHARP OBSERVABLES. *Let E be a non-trivial POV measure. There is no E -discrimination (and hence no E -measurement) $\langle \mathcal{H}_a, Z, T_a, U \rangle$ such that for all initial \mathcal{S} states T the $\mathcal{S} + \mathcal{A}$ states $U(T \otimes T_a)U^{-1}$ are mixtures of eigenstates of the projections $I \otimes Z(X_i)$ for a given reading scale (X_i) .*

PROOF: First we show that for a non-trivial E there exist pairs of orthogonal vectors which are E -distinguishable. Indeed assume that for any pair of orthogonal unit vectors $\{\varphi_1, \varphi_2\}$ one has $p_{\varphi_1}^E = p_{\varphi_2}^E$. Then also the pairs $\{\frac{1}{\sqrt{2}}(\varphi_1 + \varphi_2), \frac{1}{\sqrt{2}}(\varphi_1 - \varphi_2)\}$ and $\{\frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2), \frac{1}{\sqrt{2}}(\varphi_1 - i\varphi_2)\}$ are E -indistinguishable. From this one infers that $\langle \varphi_1 | E(X)\varphi_2 \rangle = 0$. Since φ_1, φ_2 can be arbitrary members of any orthonormal basis, it follows that $E(X)$ is diagonal in each such basis and hence $E(X) = \lambda(X)I$ for all $X \in \Sigma$. Thus E would be trivial, which was excluded.¹² Suppose φ_1, φ_2 are two orthogonal unit vectors and E -distinguishable. Suppose $U(P_{\varphi_1} \otimes T_a)U^{-1}$ and $U(P_{\varphi_2} \otimes T_a)U^{-1}$ are

$I \otimes Z$ —distinguishable and are both expressible as mixtures of eigenstates of the projections $I \otimes Z(X_i)$. Then the conditions of the inheritance of superpositions theorem are satisfied. Therefore $U(P_\varphi \otimes T_a)U^{-1}$ is not a mixture of eigenstates of the $I \otimes Z(X_i)$, if φ is the superposition $c_1\varphi_1 + c_2\varphi_2$ with nonzero coefficients c_1, c_2 . On the other hand, if vectors φ_1, φ_2 with the assumed properties do not exist, then the claim of the present theorem would also hold.

We remark as a by-product that this result encompasses all POV measures that may be introduced for other reasons than dealing with inaccurate measurements, such as joint observables for noncommuting collections of observables.

4. Conclusion.

We have extended earlier insolubility results in two ways: first, we have taken the object observable to be an arbitrary PV measure (sharp observable), and second we have taken it to be an arbitrary POV measure (unsharp observable when not a PV measure). A further step towards more realistic measurement situations would be to consider measurement schemes where the pointer observable Z itself is an unsharp observable. It is a largely open question whether an insolubility theorem can be proved in such cases.¹³

Acknowledgements.

The research of one author (AS) was supported by the National Science Foundation under grant #PHY-93-21992. This work was carried out during PB's stay at Harvard University, funded by a Feodor-Lynen-Fellowship of the Alexander von Humboldt-Foundation.

Footnotes and References.

1. E. P. Wigner, *Am. J. Phys.* **31**, 6 (1963).
2. B. d'Espagnat, *Nuovo Cim. Suppl.* **4**, 828 (1966), and *Conceptual Foundations of Quantum Mechanics*, Benjamin, Menlo Park, 1971, pp. 331-339.
3. A. Fine, *Phys. Rev. D* **2**, 2783 (1970).
4. A. Shimony, *Phys. Rev. D* **9**, 2321 (1974).
5. P. Busch, P. Lahti, P. Mittelstaedt, *The Quantum Theory of Measurement*, Springer-Verlag, Berlin, 1991. 2nd ed. in press, 1996.
6. H. Brown, *Found. Phys.* **16**, 857 (1986).
7. A. Fine, in *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, R. Kargon and P. Achinstein (eds.), MIT Press, Cambridge, MA, 1987, pp. 491-506.
8. A. Shimony, *Search for a Naturalistic World View, Vol. II*, Cambridge University Press, Cambridge, UK, 1993, Section 2, pp. 41-47.
9. H. Stein, On a Theorem in the Quantum Theory of Measurement, in *Experimental Metaphysics – Quantum Mechanical Studies in Honor of Abner Shimony*, D. Reidel, Dordrecht, to appear.
10. P. Busch, M. Grabowski, P. Lahti, *Operational Quantum Physics*, Springer-Verlag, Berlin, 1995.
11. In Ref. 10 the term *unsharp observable* is introduced in a somewhat more restrictive sense, which can be ignored for the present purpose.
12. Note that for a Hilbert space of dimension greater than 2 there is a simpler argument: for any two unit vectors which are not collinear there is at least one unit vector which is orthogonal to both. Then it is obvious that E -indistinguishability for orthogonal pairs of vectors implies E -indistinguishability for arbitrary pairs.
13. P. Busch, Can 'Unsharp Objectification' Solve the Quantum Measurement Problem?, preprint, 1996.